Combined Attacks for Cryptanalysis of Block Ciphers

Orr Dunkelman

Dept. of Electrical Engineering ESAT-SCD/COSIC
Katholieke Universitiet Leuven

Joint work with Eli Biham and Nathan Keller

December 12th, 2006

Topics of the Talk

1. Block Ciphers
2. Cryptanalysis of Block Ciphers
3. Differential Cryptanalysis
4. Linear Cryptanalysis
5. Combined Attacks
6. New Combined Attacks
7. Summary

Block Ciphers

- Basic cryptographic primitive
- Symmetric key primitive
- Can be transformed into stream ciphers, hash functions, and message authentication codes
- Many candidates for "good" block cipher:
  - Data Encryption Standard (DES)
  - Advanced Encryption Standard (AES)
  - IDEA, Blowfish, Serpent, ...

Block Ciphers (cont.)

- Pair of algorithms: encryption and decryption
- Each accepts blocks of plaintexts in a given size, e.g., 64 or 128 bits
- Encryption defines a keyed permutation of plaintexts into ciphertexts

Due to implementation issues, most block ciphers are iterated:

- Encryption is done in rounds
- Each round applies a relatively "weak" transformation to the data
- The key is used to generate set of subkeys

Generic Attacks

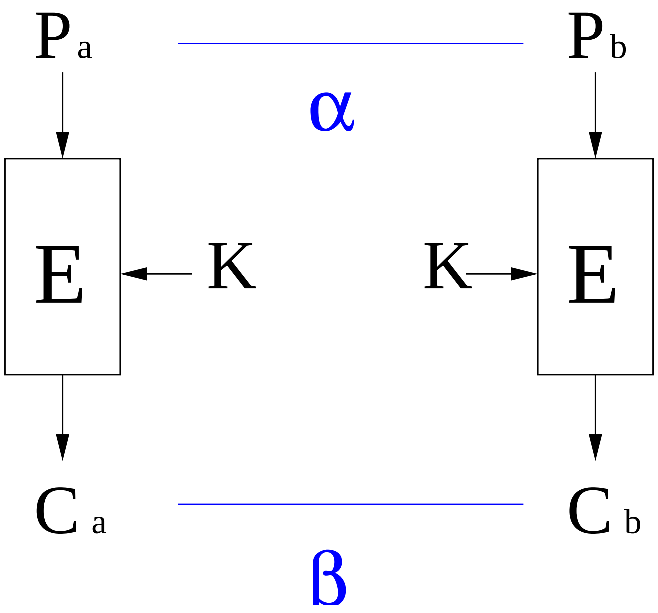
- Exhaustive key search
- Dictionary attack
- Time-memory-data trade-offs

Cryptanalysis Techniques

- Using bad mixture of data with key
- Statistical properties of the cipher
- Structural properties of the cipher
- Algebraic properties of the cipher

Differential Cryptanalysis

- Introduced by Biham and Shamir in 1990
- Studies the development of differences through the encryption function
- A differential $\alpha \to \beta$ with probability p:



- $O(p^{-1})$ pairs needed to use the differential.

Differential Cryptanalysis (cont.)

Possible generalizations:

- Truncated differentials
- Higher-Order differentials
- Impossible differentials

Linear Cryptanalysis

- Introduced by Matsui in 1993
- Studies linear approximations of the encryption function
- A linear approximation $\lambda_P \to \lambda_C$ with probability $1/2 \pm q$:

$$\lambda_P \cdot P \oplus \lambda_C \cdot C = \lambda_K \cdot K$$

for $1/2 + q$ of the plaintext-ciphertext pairs (P, C)

- The sign of the bias q is less important
- Requires about $O(q^{-2})$ known plaintexts

Distinguisher, Key Recovery, and the All the Rest

The statistical properties define *distinguishers*. These distinguishers can distinguish whether a black box is a given cipher or a random permutation.
If there is a distinguisher for r rounds of the cipher, key recovery on $r+i$ rounds is usually:

1. Guess the subkey of i rounds
2. Partially encrypt/decrypt the data the i rounds
3. Run the distinguisher
4. If the distinguisher fails — discard subkey guess
5. If the distinguisher succeeds — right subkey found*

Combined Attacks

Modern ciphers try to avoid good long statistical properties:

- Avalanche criteria
- Wide-trail strategy
- Highly nonlinear functions
- Many rounds

However, good short properties still exist, and are hard to be avoided.

Combined Attacks (cont.)

Treats the cipher E as a cascade of two sub-ciphers, i.e., $E = E_1 \circ E_0$.
Use one statistical property for the first sub-cipher, a second one for the second sub-cipher, and stitch them somehow.

- First example — differential-linear (Langford and Hellman, 1994)
- Miss in the middle attacks (Biryukov, Biham, Shamir, 1999)
- Boomerang attack (Wagner, 1999)

Differential-Linear Attack

- Use differential $\Omega_P \to \Omega_T$ for E_0 (with probability 1)
- Use linear approximation $\lambda_T \to \lambda_C$ for E_1
- Ask for the encryption of the pair $(P_1, P_2 = P_1 \oplus \Omega_P)$
- The intermediate encryption values satisfy $T_1 \oplus T_2 = \Omega_T$
- The value difference between $\lambda_T \cdot T_1$ and $\lambda_T \cdot T_2$ is known (equals to $\lambda_T \cdot \Omega_T$)
- The values $\lambda_C \cdot C_1$ and $\lambda_C \cdot C_2$ are equal with probability $1/2 \pm 2q^2$

Miss-in-the-Middle Attack

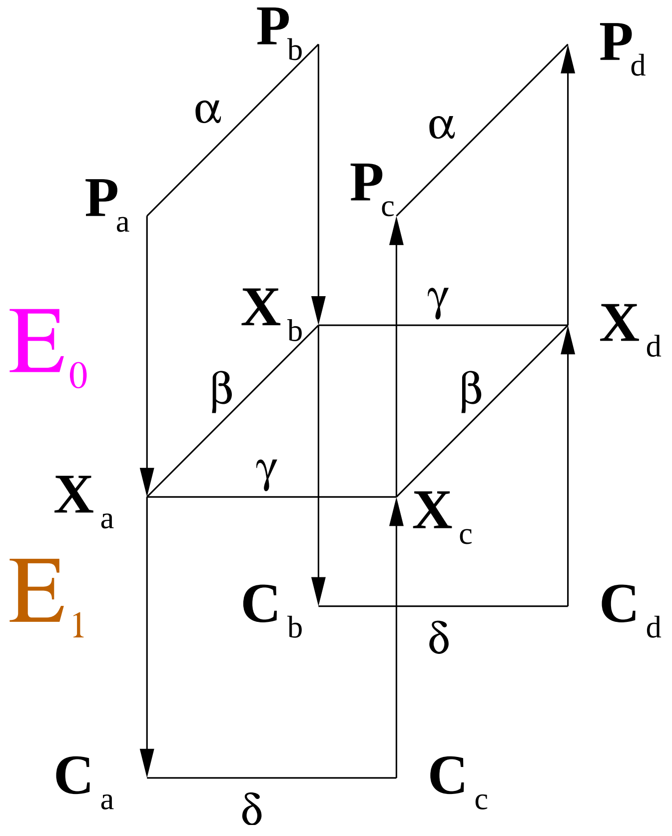
- Use (truncated) differential $\alpha \to \beta$ for E_0
- Use (truncated) differential $\delta \to \gamma$ for E_1^{-1}
- If β and γ cannot be the same, e.g., first bit of β is always active, while first bit of γ is zero, then
$$\alpha \not\to \delta$$
- Any key suggesting that $\alpha \to \delta$ is obviously wrong

Note: The concept of using impossible events to discard wrong key guesses was used in the attacks on the Enigma.

The Boomerang Attack

Uses two differentials, $\alpha \longrightarrow \beta$ for E_0, and $\gamma \longrightarrow \delta$ for E_1.

The structure of the boomerang attack:



Probability of obtaining right quartet p^2q^2.

The Boomerang Attack (cont.)

There are several possible improvements:

1. Use several $\gamma' \longrightarrow \delta$ differentials.
2. Use several $\alpha \longrightarrow \beta'$ differentials.
3. Use different differentials for each pair:

$$\begin{array}{l} \alpha \longrightarrow \beta \text{ for } (P_1, P_2) \\ \gamma \longrightarrow \delta \text{ for } (C_1, C_3) \\ \gamma \oplus A \longrightarrow \delta \text{ for } (C_2, C_4) \\ \alpha \longrightarrow \beta \oplus A \text{ for } (P_3, P_4) \end{array}$$

The overall probability is

$$\sum_{\beta,\gamma,A} \Pr[\alpha \longrightarrow \beta] \cdot \Pr[\alpha \longrightarrow \beta \oplus A] \cdot \Pr[\gamma \longrightarrow \delta] \cdot \Pr[\gamma \oplus A \longrightarrow \delta]$$

The Boomerang Attack (cont.)

Computing the exact probability is hard, but it can approximated using only the first two improvements. This yield:

$$\sum_{\beta',\gamma'} \Pr{}^2[\alpha \to \beta'] \cdot \Pr{}^2[\gamma' \to \delta] = \hat{p}^2\hat{q}^2$$

New Combined Attacks

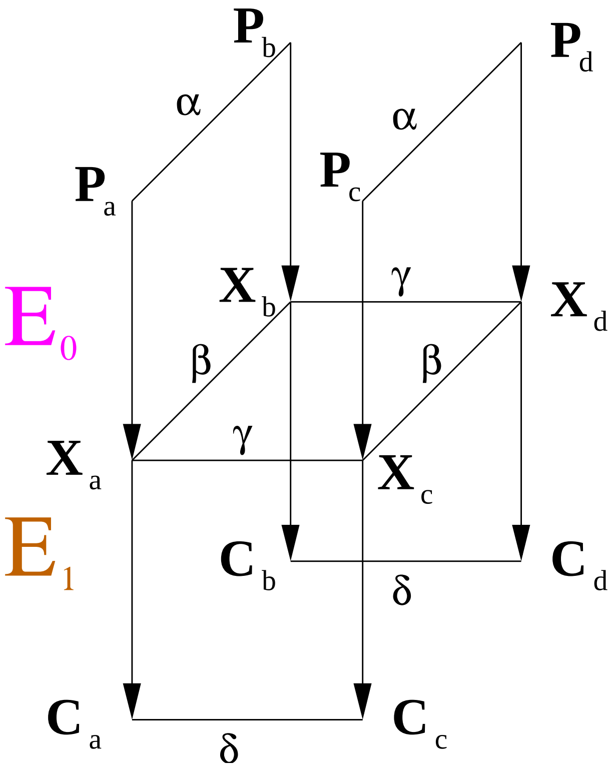
- Amplified Boomerang/Rectangle
- Enhanced Differential-Linear
- Differential-Bilinear
- Higher-Order Differential-Linear
- (Differential-)(Bi)Linear Boomerang

The Amplified Boomerang

Introduced by Kelsey, Kohno, and Schneier in 2000.
Transforms the boomerang attack into a chosen plaintext attack.

- Transformation is based on birthday paradox arguments.
- Right quartet is formed with probability $2^{-(n+1)/2}pq$.
- New problem — identifying the right quartets.

The Amplified Boomerang (cont.)



Probability of obtaining right quartet $2^{-n-1}p^2q^2$.

Due to the birthday paradox $O(2^{(n+1)/2}/pq)$ pairs suffice for obtaining a right quartet.

The Rectangle Attack

Improves the amplified boomerang attack:

- Better probability for right quartet, using multiple differentials for each sub-cipher (just like in the boomerang attack) — $\hat{p}^2\hat{q}^2$.
- Better usage of the data (twice the possible quartets for the same data)
- Better algorithms for finding the right quartets

Enhanced Differential-Linear Attack

Allows using differentials with probability less than 1.

Assume that the differential $\Omega_P \to \Omega_T$ has probability p.

The probability that $\lambda_T \cdot T_1 = \lambda_T \cdot T_2$ is

$$\underbrace{p}_{\text{differential holds}} + \underbrace{(1-p)\cdot\frac{1}{2}}_{\text{differential doesn't hold}} = \frac{1}{2} + \frac{p}{2}$$

The probability that $\lambda_C \cdot C_1 = \lambda_C \cdot C_2$ is

$$\frac{1+p}{2}\cdot\left(\frac{1}{2}+2q^2\right) + \frac{1-p}{2}\cdot\left(\frac{1}{2}-2q^2\right) = \frac{1}{2}+2pq^2$$

Can be used also when $\Omega_T \cdot \lambda_T$ is unknown (but constant).

Differential-Bilinear Attack

Bilinear cryptanalysis is a generalization of linear attack to Feistel block ciphers, presented by Courtois in 2004.

$$L_0[\alpha_0] \cdot R_0[\beta_0] \oplus R_0[\gamma_0] \oplus L_0[\delta_0] \oplus L_n[\alpha_n] \cdot R_n[\beta_n] \oplus R_n[\gamma_n] \oplus L_n[\delta_n] =$$

$$L_0[\epsilon_0] \cdot K[\epsilon_1] \oplus R_0[\zeta_0] \cdot K[\zeta_1] \oplus L_n[\eta_0] \cdot K[\eta_1] \oplus R_n[\theta_0] \cdot K[\theta_1] \oplus K[\iota_1]$$

Differential-Bilinear Attack (cont.)

The attacker chooses pairs of plaintexts P_1 and $P_2 = P_1 \oplus \Omega_P$.
With probability p the intermediate encryption values T_1 and T_2, have a difference that satisfies the equality

$$T_{1L}[\alpha_0] \cdot T_{1R}[\beta_0] \oplus T_{1L}[\gamma_0] \oplus T_{1R}[\delta_0] = T_{2L}[\alpha_0] \cdot T_{2R}[\beta_0] \oplus T_{1L}[\gamma_0] \oplus T_{2R}[\delta_0],$$

Similarly to the differential-linear case, the pair of ciphertexts C_1 and C_2 satisfies the following equation

$$C_{1L}[\alpha_n] \cdot C_{1L}[\beta_n] \oplus C_{1L}[\gamma_n] \oplus C_{1R}[\delta_n] = C_{2L}[\alpha_n] \cdot C_{2R}[\beta_n] \oplus C_{2L}[\gamma_n] \oplus C_{2R}[\delta_n]$$

with probability $1/2 + 2pq^2$.

Differential-Bilinear Attack (cont.)

There are several differences between differential-linear and differential-bilinear attacks:

- Not any differential can be used in the differential-bilinear attack
- Differential must give the attacker the difference of $L_T[\alpha_0] \cdot R_T[\beta_0]$
- Subkey terms exist in the bilinear approximation

Higher-Order Differential-Linear Attack

Higher-order differentials are differentials of sets of plaintexts.

The common higher-order differential predicts that the XOR value of all ciphertexts is some constant (usually zero).

Our proposed attack can use also SQUARE-like properties (SQUARE, saturation, and Integral attacks).

All we require is that the XOR of some bit(s) over all the intermediate encryption values is fixed.

Higher-Order Differential-Linear Attack (cont.)

Let *Set* be a set of plaintexts $\{P_1, P_2, \ldots, P_m\}$ such that the higher-order differential predicts (with some probability p) the value

$$\oplus_{i=1}^{m} T_i.$$

Let the event I be

$$I = \{\lambda_P \cdot (T_1 \oplus \ldots \oplus T_m) = \lambda_C \cdot (C_1 \oplus \ldots \oplus C_m)\}.$$

Then (under standard independence assumptions) $\Pr[I] = 1/2 + 2^{m-1} q^m$. This leads to bias of $2^{m-1} p q^m$ in the value of

$$\oplus_{i=1}^{m} \lambda_C \cdot C_i$$

Higher-Order Differential-Linear Attack (cont.)

We can use the higher-order differential-linear attack for:

- Transform any linear approximation with bias 1/2 for r Feistel rounds into an higher-order differential-linear approximation with bias 1/2 for $r + 3$ rounds
- When m small, or q very high (very common for weak key classes)

We note that the higher-order differential-linear attack was developed independently by Shin et al. under the name square-nonlinear attack.

(Differential-)(Bi)Linear Boomerang Attack

Our main observation is that the boomerang process leads to a pair of intermediate encryption values with difference β.
This β difference may be used in (differential-)(bi)linear attack (on the first sub-cipher).

- The cipher is actually decomposed into three sub-ciphers
- If the probability of the first differential p_0, of the second differential p_1, and the (differential-)(bi)linear approximation of E_0 has bias $(2p)q$, the bias for the entire structure is:

$$2p_0\hat{p}_1{}^2(p)q$$

Results of the Rectangle Attack

Cipher	Number of Rounds	Complexity Data	Complexity Time	Number of Keys
Serpent	10	$2^{126.3}$ CP	$2^{173.8}$	1
SC2000	3.5	$2^{83.1}$ CP	$2^{83.1}$	1
SHACAL-1	49	$2^{151.9}$ CP	$2^{508.4}$	1
	full	$2^{159.8}$ RK-CP	2^{420}	4
	full	$2^{153.8}$ RK-CP	$2^{501.2}$	4
AES-192	9	2^{86} RK-CP	2^{181}	32
AES-256	10	$2^{114.9}$ RK-CP	$2^{199.8}$	32
COCONUT98	full†	4 RK-ACPC	1	2
IDEA	5.5†	$2^{51.6}$ RK-ACPC	1	4
	6.5	$2^{59.8}$ RK-CP	$2^{88.1}$	4
	7	2^{65} RK-CP	$2^{104.2}$	4
KASUMI	full	$2^{54.6}$ RK-CP	$2^{91.7}$	4

† – Distinguishing attack, RK – Related-key, CP – Chosen plaintext, ACPC – Adaptive chosen plaintext and ciphertext

Results of the enhanced differential-linear attack

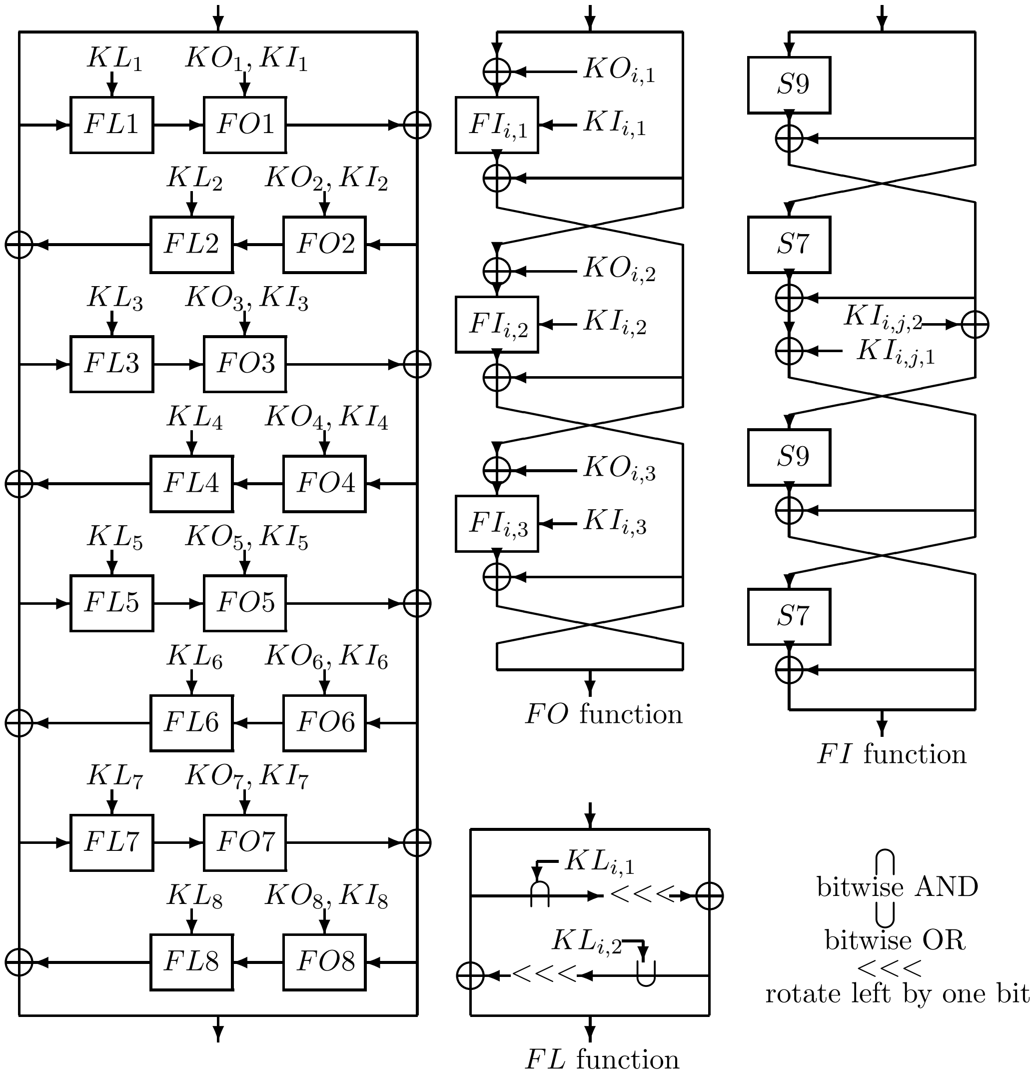
Cipher	Number of Rounds	Complexity Data	Complexity Time
DES	8	$2^{14.8}$ CP	$2^{14.8}$
	9	$2^{15.8}$ CP	$2^{29.2}$
COCONUT98	full	$2^{27.7}$CP	$2^{33.7}$
Serpent	11	2^{118}KP	$2^{205.7}$

The Case for KASUMI

- KASUMI is the 3GPP block cipher
- 8-round Feistel; 64-bit block; 128-bit key
- Used to ensure privacy and authenticity of the conversations
- The mode of use are proved to be secure, if the cipher is secure against related-key attacks

Our attack on KASUMI is a related-key rectangle attack, which is a related-key attack...

The Case for KASUMI (cont.)



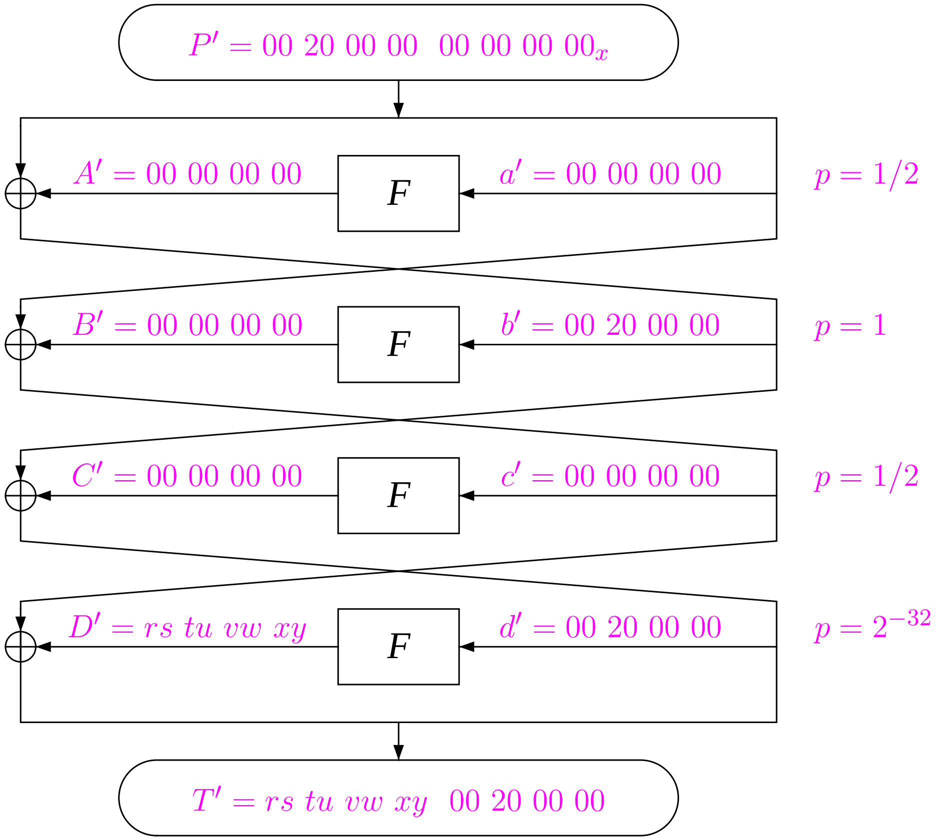
The Case for KASUMI (cont.)

Round	$KL_{i,1}$	$KL_{i,2}$	$KO_{i,1}$	$KO_{i,2}$	$KO_{i,3}$	$KI_{i,1}$	$KI_{i,2}$	$KI_{i,3}$
1	$K_1 \lll 1$	K'_3	$K_2 \lll 5$	$K_6 \lll 8$	$K_7 \lll 13$	K'_5	K'_4	K'_8
2	$K_2 \lll 1$	K'_4	$K_3 \lll 5$	$K_7 \lll 8$	$K_8 \lll 13$	K'_6	K'_5	K'_1
3	$K_3 \lll 1$	K'_5	$K_4 \lll 5$	$K_8 \lll 8$	$K_1 \lll 13$	K'_7	K'_6	K'_2
4	$K_4 \lll 1$	K'_6	$K_5 \lll 5$	$K_1 \lll 8$	$K_2 \lll 13$	K'_8	K'_7	K'_3
5	$K_5 \lll 1$	K'_7	$K_6 \lll 5$	$K_2 \lll 8$	$K_3 \lll 13$	K'_1	K'_8	K'_4
6	$K_6 \lll 1$	K'_8	$K_7 \lll 5$	$K_3 \lll 8$	$K_4 \lll 13$	K'_2	K'_1	K'_5
7	$K_7 \lll 1$	K'_1	$K_8 \lll 5$	$K_4 \lll 8$	$K_5 \lll 13$	K'_3	K'_2	K'_6
8	$K_8 \lll 1$	K'_2	$K_1 \lll 5$	$K_5 \lll 8$	$K_6 \lll 13$	K'_4	K'_3	K'_7

$X \lll i$ — X rotated to the left by i bits

The Case for KASUMI (cont.)

Related-Key Differential Characteristic of Rounds 1–4:



Key difference $\Delta K_{ab} = (0, 0, 1_x, 0, 0, 0, 0, 0)$.

The Case for KASUMI (cont.)

There are 2^{32} possible values for $(rs\ tu\ vw\ xy)$, so counting all differentials we obtain:

$$\hat{p}^2 = \sum_{rs\ tu\ vw\ xy} \Pr{}^2[P' \to T'] = \sum_{rs\ tu\ vw\ xy} 2^{-34\cdot 2} = 2^{32} \cdot 2^{-68} = 2^{-36}$$

- A small improvement, fixing two plaintext bits, increases the probability by a facotr of 2 ($\hat{p}^2 = 2^{-34}$).
- The differential for rounds 5–7 is the first three rounds of the above differential ($\hat{q}^2 = 2^{-4}$).

The Attack Algorithm

1. Take 2^{51} pairs of plaintexts and ask for their encryption under two pairs of keys. ($\Rightarrow 2^{102}$ quartets)

2. Look for quartets whose pairs have difference 00 20 00 00 in the right half. ($\Rightarrow 2^{38}$ quartets remain)

3. For each quartet, guess $KO_{8,1}$ and $KI_{8,1}$, and obtain the input difference to FL_8 in 16 bits (for the two pairs). Deduce the value of $KL_{8,2}$. About 2^{-16} of the (quartet,guesses) remain ($\Rightarrow 2^{38+32-16} = 2^{54}$ (quartets,subkey guesses) tuples).

4. Repeat the above with $KO_{8,3}, KI_{8,3}$, and obtain $KL_{8,1}$. ($\Rightarrow 2^{54+32-16} = 2^{70}$ (quartets,subkey guess) tuples).

5. Try the remaining 32 bits of the key.

Questions?

Thank You!